PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

Eric HELLMAN et al.

Art Unit: Unknown

Application No.: 09/437,378

Examiner: Unknown

Filed: November 10, 1999

For: RETRIEVAL OF DIGITAL OBJECTS BY REDIRECTION OF CONTROLLED VOCABULARY SEARCHES

DECLARATION IN SUPPORT OF PROTEST UNDER 37 C.F.R. §1.291

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

August 1, 2003

Sir:

The following declaration makes reference to several supporting documents that are included herewith, either in the included PTO-1449 or as attachments to this Declaration.

I, HERBERT VAN DE SOMPEL, declare that:

1. I was born in Ghent, Belgium, on March 20th 1957. In 1979, I received a Masters in Mathematics, and in 1981 a Masters in Computer Science, both from Ghent University. In June 2000, I received a PhD in Communication Science from Ghent University with a thesis entitled "Dynamic and Context-Sensitive Linking of Scholarly

Information". The thesis is a compilation of previous publications in which I describe a novel way to link scholarly information on the Web, which would later become known as the OpenURL Framework (for Context-Sensitive Services). My professional career has been fully focused on the intersection of scholarly information and information technology. For 17 years, I was Head of Library Automation at Ghent University. After having left Ghent in 2000, I have been Visiting Professor in Computer Science at Cornell University, and Director of e-Strategy and Programmes at the British Library. Currently, I am the team leader of the Digital Library Research and Prototyping Team at the Research Library of the Los Alamos National Laboratory.

2. In 1998, I was still heading Library Automation at Ghent University in Belgium. Like every academic library, we used Web-technologies to provide students and researchers with access to many different information resources, stored both on our University Intranet and remotely on the Web. Some information resources were in open access, but most were accessible on a subscription-basis only. Over time, every information vendor we dealt with had started to implement its proprietary approach to link to information from other vendors. But the links they provided frustrated us, and other libraries worldwide:

- The links failed to take into account the environment of the user following it: vendor A would present a link into the collection of vendor B, irrespective of whether the user (or the user's institution) had a subscription to the collection of vendor B. If the user didn't have one, she was confronted with a frustrating HTTP 404 error. Also, typically an institution would subscribe to the collection of

vendor B through one of many intermediaries, had to access that collection via the Web site of exactly that intermediary, and could not gain access to the collection through another intermediary. Unfortunately, the links provided by vendor A would not take this reality into account, and would just bring the user to what the vendor considered to be the "default" access point to the collection of vendor B. Again, a frustrating HTTP 404 resulted when the user's access point was not the one vendor A had decided to be the "default" one.

- The links were limited in 2 ways:
    - Typically vendor A would only provide links to the collection of vendor B, if A and B had some business agreement. As a result, many collections that could (and from the library's perspective should) be interlinked were in fact not.
    - Vendor A would unilaterally decide on the kind of service a link would provide. For example, the vendor would decide it should go from an abstract in its own collection that described a scholarly paper to the electronic full-text at the site of the publisher that had published the paper. Libraries wanted other kinds of links, for examples, links to check whether they had a paper copy of the article in their stacks.

3. In the Spring of 1998, I conceived of a solution to the above problem. Instead of having vendor A provide a link into the collection of vendor B, I thought that vendor A should provide a link to a, then non-existing, Web-based service that each library could operate. When a user would click that link, she would actually transport information

from the collection of vendor A to the new Web-based service operated by her library. The transported information (metadata and identifiers) would obviously pertain to the information resource for which the link was provided; in the above example that resource is the record containing the abstract of the paper. The library's Web-based service would analyze the transported information, and, based on contained knowledge (business rules and information about the library collection) provide a menu of links to the user. Those links would not lead to the frustration described above, because the library administration could tailor them to match the local situation.

4. I demonstrated an early prototype illustrating these ideas before a large audience at a meeting of the Flemish Library Association VOWB (see http://www.libis.be/vowb/) in October 1998. In August 1998, I had already briefly described the solution in a lecture at the TICER Summer School on Digital Libraries (see http://www.ticer.nl/summer98/). In April 1999, I published 2 papers in the open access Web journal for digital library research D-Lib Magazine (http://www.dlib.org). One paper described the problem statement, the other the nature of my proposed solution and a description of the aforementioned prototype. Early 1999, I moved to the Los Alamos National Laboratory to do further experiments to demonstrate the feasibility of my solution. In those elaborate experiments, major information providers such as the American Physical Society, Wiley Publishers, the arxiv.org Physics server, Ex Libris, SilverPlatter participated. Both the American Physical Society and Wiley published a press release referencing the experiments (see http://www.aps.org/apsnews/0899/089914.html and

http://www.library.yale.edu/~llicense/ListArchives/9907/msg00012.html). In June of 1999, I described the nature of my solution in front of a group of experts (including an Applicant of the above-cited application, "Mr. Eric Hellman") at the NISO Reference Linking meeting in Boston (report of the meeting, including my slides at http://www.niso.org/news/events_workshops/linkrept.html). Later that month, I described the solution and demonstrated results from the Los Alamos experiments at a session of the Library Technology Association (LiTA) held at the occasion of the Summer Meeting of the American Library Association in New Orleans. Many considered my presentation to be the highlight of that Meeting. In October 1999, I published on the evolution of my linking research in D-Lib Magazine; the paper includes a detailed description of the general architecture, of the individual components therein, and the results of the Los Alamos experiments. More specifically, the paper details:

- A controlled vocabulary for the link from a vendor to the library's Web-based service. The controlled vocabulary was named the SFX-URL. The SFX-URL transported identifiers and metadata about the resource for which the link was provided from the vendor to the library's Web-based service.
- The nature and structure of the library's Web-based service. This Web-based service was named the service component, and our implementation was named the SFX server.
- A technique to allow the vendor to point the controlled vocabulary link to the Web-based service of a specific user. The technique was named the CookiePusher.

5. Starting June 1999, Mr. Eric Hellman started to lobby me to file an application with him to patent the concepts from my linking research. I have archived several e-mails to illustrate that discussion, where copies of the emails have been attached hereto. After I decided not to file with Mr. Eric Hellman, he apparently decided to proceed by himself.

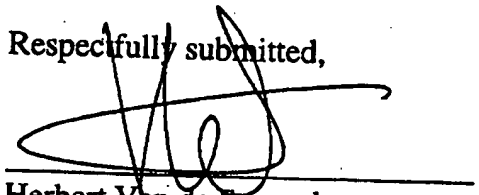
6. Early 2000, Ex Libris acquired the SFX software from Ghent University (see http://www.sfxit.com/). In May 2000, I published on the Web the OpenURL syntax describing the controlled vocabulary that could be used for the link from the vendor to the library's Web-based services (see http://www.sfxit.com/openurl/openurl.html). Very rapidly, information vendors started implementing it. In June 2000, Ex Libris began testing SFX in several US libraries, and later on began marketing it as a product. Since then many vendors of library software have started marketing their own version of the library's Web-based service, now commonly called an OpenURL linking server, OpenURL service component or OpenURL resolver. On December12, 2000 I and Oren Beit-Arie submitted a proposal to the National Information Standards Organization to standardize the OpenURL syntax as a NISO and ANSI American National Standard. In preparation of the standardization process that started in June 2001, we published a paper in D-Lib Magazine, which generalized my linking concepts from the scholarly information sub-domain of the Web, to the Web in general.

7. I have reviewed a version of the above-cited application as was published by the World Intellectual Property Organization as WO 01/35279 on May 17, 2001 and I believe that it largely describes my work, as evidenced by the references cited in this

Declaration.

8. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Respectfully submitted,



Herbert Van de Sompel

Enclosures: Copies of e-mail communication between the Declarant and one of the applicants of the above-cited application

Subject: Re: Baton-Go
Date: Wed, 16 Jun 1999 20:47:45 -0400
From: Eric Hellman <eric@hellman.net>
To: herbert van de sompel <hvds@lanl.gov>

I talked to a patent attorney today. He expects that getting a patent will cost ~$10K.

At this point, we need to start doing a few things.

First, before we try to obtain a patent, we should decide on an arrangement about who is to own the patent, who pays what, etc. Think about it. I can front the money, but you need to decide how to protect your interests.

Next, we need to assemble documentation about exactly what was invented and when it was invented. In the US, it matters when something was invented; in other countries, it's date of filing.

In my mind the invention consists of 3 concepts.
1. Controlled search vocabulary requests
2. a central database of search service configurations
3. the "baton" concept of passing the request along

If you can add the certificate concept that would be good.

I'm off to NIH tomorrow.

Eric

Eric Hellman, President
eric@openly.com
tel/fax 1-973-509-7800
Tools for 21st Century Scholarly Publishing

Openly Informatics, Inc.
10 Columbus Ave., Suite C
Montclair, NJ 07042
http://www.openly.com/

Printed on 2003/07/31

Subject: Baton-Go
Date: Mon, 28 Jun 1999 23:55:13 -0400
From: Eric Hellman <eric@hellman.net>
To: herbert van de sompel <hvds@lanl.gov>

Heard you made a nice splash at ALA.

Here's a draft of a disclosure.

I propose we deal with ownership as follows:
1. If it's not big, it's not worth worrying about.
2. Openly will pay all expenses related to obtaining patent protection. Openly will in return have exclusive rights for 2 years.
3. At the end of 2 years, you will have the option to buy the rights to the technology for $100,000. Openly can then choose to either (a) sell you the rights. or (b) pay you $100,000 for your interest in the rights.

I am working on a business plan. I've discussed it with a number of investors, and there is a fair amount of interest. Included among the people who are interested are Eugene Garfield, Founder of ISI, Jim Carlisle, EVP of Think New Ideas, and Vitek Tracz, CEO of Current Science Publishers.

Eric

disclosure.doc	**Name:** disclosure.doc **Type:** WINWORD File (application/msword) **Encoding:** base64

Eric Hellman, President
eric@openly.com
tel/fax 1-973-509-7800
Tools for 21st Century Scholarly Publishing

Openly Informatics, Inc.
10 Columbus Ave., Suite C
Montclair, NJ 07042
http://www.openly.com/

Printed on 2003/07/31

PRINTOUT OF DISCLOSURE.DOC 3 pages
Printed on 2003/07/31

System for Retrieval of Digital Objects by Redirection of Controlled-Vocabulary Searches

Abstract

We disclose a system which enables the construction of hyperlinks from an open information network such as the World-wide web which target information in restricted-access information services. A controlled vocabulary is used to form abstract hyperlinks, which are made concrete by an intermediary with access to user prefences.

Background

Although the World-Wide Web has grown tremendously over the last 5 years, there are vast stores of information which are still inaccessible. Books, popular literature, scholarly journals, and subscription-based information services are starting to be delivered over the Internet, but because of access restrictions, they have benefitted only marginally from the navigational infrastructure of the World-wide-web. Search engines cannot access subscription-based content, and users of such content are deprived of the benefits of hyperlinks from content outside the subscription based service. For these and similar reasons, revenue from subscription based services has been dwarfed by the financial success of free-to-access services such as Yahoo. At the same time, the availability of free-to-access information has threatened the economic viability of higher-quality paid-access and restricted-access services.

In this disclosure, we teach a new system for the construction and resolution of hyperlinks which has the potential to integrate restricted-access information services such as digital libraries into the larger world-wide web.

Hyperlinks

Hypertext was invented by Nelson [Nelson] in the 60's as a vision for how computers could fundamentally change the way humans use text. Hypertext used hyperlinks to allow a user to jump from one point in a text to another, or from one text to a completely different text. Today's world-wide web is based on a hypertext format called HTML, in which hyperlinks are embedded into text using special tags and standard addresses called URL's. The URL might point to another location in the same document, to another document from the same server computer, or to a point in another document from a server computer anywhere in the world. Graphic images are embedded into documents in a similar way; the URL for a graphic could point at a graphic file located anywhere in the world.

Hyperlinks assist a user's navigation through the network of information. They are put in by an author based on guesses about what a reader might be interested. The reader then follows a path of his own choosing based on his own interests. The process of authoring hyperlinks can be demanding and laborious. In an open environment, the hyperlinks must also be

maintained based on the changing environment. Hyperlink authors typically know little about the resources that users have access to, and by necessity, they traditionally can only make links to resources that exist at the time of authoring.

Link Abstraction

One technique known in the art for simplifying the process of hyperlink authoring is to use abstract links, or search query links. For example, a hyperlink could be added for the words "Chester A. Arthur" by embedding a URL which specifes an appropriate search to a search engine available to the readers of the hypertext. For example, to send this search to the "Alta Vista Search Engine, one could use the URL http://www.altavista.com/search?term=Chester+A.+Arthur Since this search is done at the dime the reader follows the link, the search can return documents that did not exist at the time that the hyperlink was authored. Minor et al. taught that search queries could also contain "demographic information" about a user which would allow a search engine to customize its response to the link. This approach required the linking site to identify and collect data about the users of the linking site, which is true for many commercial applications.

Another technique of link abstraction. known in the art is to use object names rather than addresses in the links. A name resolver is then used by the reader to resolve such a link. One advantage of this approach is that the user can be directed to a local copy of the object named in the link. Efforts to implement name resolvers on the internet have included the URN, the PURL, and the DOI. These efforts have met with limited market acceptance and deployment.

A widely used technique for information discovery is the use of metadata for categorizing information. Metadata is the use of a controlled vocabulary, or schema, to describe information items. For example, the widely used "Dublin Core" vocabulary defines specific property names, such as "publisher" or "creator" to allow particular means of accessing large numbers of data objects.

A technique known to those skilled in the art for abstraction of links is controlled vocabulary link construction. The values of controlled vocabulary words describing a particular data object can be assembled in a specified ways using mathematical formulae and algorithms to form URL's which can be embedded in web pages to form hyperlinks for users.[S-Link-S]

Link Indirection

Another type of technique known in the art to enhance hyperlinks is Link Indirection. In this technique, a link is made which points to an intermediary service which then redirects to user to a final destination. A good example of this are advertising banner brokers such as DoubleClick. The embedded graphics are addressed by URL's which point at the DoubleClick intermediary service. This service delivers advertising graphics which may depend on the user's past behavior, the time of day,

the referring page, or other factors, and manages the flow of payments from advertisers to content providers.

Linking into restricted services

In the example cited above, it is quite likely that the specified search would fail to inform the user that Chester A. Arthur was President of the United States. More specialized information services may deliver much better information, but are less likely to attract the wide audience which makes a big web site attractive to advertisers, and thus must charge for access. By combining techniques known to those skilled in the art, we teach how a hyperlink can be made an intermediary service which resolves to the specialized information for a scholar with access to the specialized service, and which resolve to a generic service for someone without special access priveleges. This new capability benefits the reader, the hyperlink author, the provider of the specialized service and the general reader by enabling new revenue sources for the participating services. Figure 1 shows a preferred embodiment of how the enhanced hyperlink can be accomplished. A publisher embeds a special link, referred to here as the "searchbaton" link which specifies a biographical search for "Chester A. Arthur". The controlled vocabulary here uses the "genre" "biography" and person-name "Chester A. Arthur". The link points at an intemediary service, referred to here as the "searchbaton server". A user requests the resource in which the searchbaton link is embedded, and curious who Chester A. Arthur was, the user clicks on the searchbaton link. The user's computer submits the link request to the searchbaton server, which reads a "cookie" presented by the users browser. The cookie allows the searchbaton server to identify the user as having access to a private biography database. The search baton server translates the request into a form understood by the user's biography search engine, and redirect the request to the private database. If the user has not registered search preferences previously, then the user is given a chance to modify his resolution preferences, which might include resolution by a free service. The important steps here are that

1. The search request has been specified using a controlled search vocabulary.
2. The search request has been delegated to a third party intermediary.
3. The search request has been combined with user preferences to form a search request resolvable by a service to which the user has priveleged access.
4.

Glossary of Terms used in U.S. Appln. No. 09/437,378:

Appropriate copy problem -
The problem of resolving a network-based reference in different ways depending on the context of the request; also called the Harvard Problem.

Controlled vocabulary abstract link -
A well-defined syntax for encoding semantic data for transmission. Examples include: DOI (Digital object identifier); OpenURL; hyperlink; any HTTP message.

Redirection facility -
A web-based service which interprets and transmits a request for a given resource to a service provider.

Resolution service -
A web-based service at which a request for services is targeted; also called a Resolver. A Resolver is the target of the OpenURL.

User information -
Information about an authenticated end user, may include information about their organizational affiliation and/or system environment.

User preference -
Functional and/or performance and/or operational preferences of the end-user; may refer to an organization or an individual.